

SEQUENCE LISTING

<110> Pharmexa A/S

<120> PURIFICATION OF HER-2 VARIANTS

<130> P1020DK00

<160> 2

<170> PatentIn version 3.2

<210> 1

<211> 5661

<212> DNA

<213> Artificial sequence

<220>

<223> Recombinant expression plasmid derived from pMT

<220>

<221> polyA_signal

<222> (263)..(268)

<223> SV40 late polyadenylation site

<220>

<221> misc_feature

<222> (1579)..(2439)

<223> Ampicillin resistance gene, encoded by complementary strand

<220>

<221> promoter

<222> (3050)..(3415)

<223> Metallothionein promoter

<220>

<221> RBS

<222> (3493)..(3501)

<223> Kozak-like sequence

<220>

<221> CDS

<222> (3502)..(5592)

<223> DNA encoding immunogenic, his-tagged variant of human HER-2

<220>

<221> sig_peptide

<222> (3502)..(3555)

<223> BiP signal sequence

<220>

<221> misc_feature

<222> (3556)..(3597)

<223> Histidine tag

<220>

<221> mat_peptide

<222> (3556)..(5589)

```

<220>
<221> misc_feature
<222> (3598)..(3603)
<223> Dipeptidase stop sequence

<220>
<221> misc_feature
<222> (3604)..(5589)
<223> Gene coding for the hHER2MA5-5DUH protein

<220>
<221> misc_feature
<222> (4357)..(4401)
<223> Diphtheria toxoid P2 epitope

<220>
<221> misc_feature
<222> (5500)..(5562)
<223> Diphtheria toxoid P30 epitope

<400> 1
ggccgctcga gtctagaggg cccttcgaag gtaaggctat ccctaaccct ctccctcggtc      60
tcgattctac gcgtaccggt catcatcacc atcaccattg agtttaaacc cgctgatcag      120
cctcgactgt gccttctaag atccagacat gataagatac attgatgagt ttggacaaac      180
cacaactaga atgcagtgaa aaaaatgctt tatttgtgaa atttgtgatg ctattgctt      240
atttgtaacc attataagct gcaataaaca agttaacaac aacaattgca ttcattttat      300
gtttcaggtt cagggggagg tgtggaggt tttttaaagc aagtaaaacc tctacaaatg      360
tggtatggct gattatgatc agtcgacctg cagggcatgca agcttggcgt aatcatggtc      420
atagctgttt cctgtgtgaa attgttatcc gctcacaatt ccacacaca tacgagccgg      480
aagcataaaag tgtaaagcct ggggtgccta atgagtgagc taactcacat taattgcgtt      540
gcgctcaactg cccgcttcc agtcggaaa cctgtcgtgc cagctgcatt aatgaatcgg      600
ccaacgcgcg gggagaggcg gtttgcgtat tggcgctct tccgcttcct cgctcaactga      660
ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat      720
acggttatcc acagaatcag gggataacgc agggaaagaac atgtgagcaa aaggccagca      780
aaaggccagg aaccgtaaaa aggccgcgtt gctggcggtt ttccataggc tccgcccccc      840
tgacgagcat cacaaaaatc gacgctcaag tcagagggtgg cgaaacccga caggactata      900
aaagataccag gcgtttcccc ctggaaagctc cctcgtgcgc tctcctgttc cgaccctgcc      960
gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgctt ctcatacgctc      1020
acgctgttagg tatctcagtt cggtgttaggt cgttcgctcc aagctggct gtgtgcacga      1080
accccccgtt cagccccgacc gctgcgcctt atccggtaac tatcgtctt agtccaaccc      1140
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcag      1200

```

gtatgttaggc ggtgctacag agttcttcaa gtggggcct aactacggct acactagaag	1260
gacagtattt ggtatctgcg ctctgtcaa gccagttacc ttcggaaaaa gagttggtag	1320
ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtgg tttttgttt gcaagcagca	1380
gattacgcgc agaaaaaaag gatctcaaga agatccttg atctttcta cggggctga	1440
cgctcagtgg aacgaaaact cacgttaagg gatttggc atgagattat caaaaaggat	1500
cttcacccatg atcctttaa attaaaaatg aagtttaaa tcaatctaaa gtatatatga	1560
gtaaacttgg tctgacagtt accaatgctt aatcagttag gcacccatct cagcgatctg	1620
tctatttcgt tcatccatag ttgcctgact ccccgctgt tagataacta cgatacggga	1680
gggcttacca tctggcccca gtgctgcaat gataccgcga gaccacgct caccggctcc	1740
agatttatca gcaataaaacc agccagccgg aaggccgag cgccagaatg gtcctgcaac	1800
tttattccgc tccatccagt ctattaattt tgccggaa gctagagtaa gtagttcgcc	1860
agttaatagt ttgcgcaacg ttgttgccat tgctacaggc atcgtgggtg caccgtcg	1920
gtttggatg gtttcattca gtcgggttc ccaacgatca aggcgagtt catgatcccc	1980
catgttgc aaaaaagcgg ttagtcctt cggcctccg atcgttgtaa gaagtaagtt	2040
ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctt ctgtcatgcc	2100
atccgtaaga tgctttctg tgactggta gtactcaacc aagtcattct gagaatagt	2160
tatgcggcga ccgagttgct cttgcggcgtc gtcaatacgg gataataccg cgccacatag	2220
cagaacttta aaagtgcata tcattggaaa acgttcttcg gggcgaaaac tctcaaggat	2280
cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaaact gatttcagc	2340
atctttact ttaccaggcg tttctgggtg agcaaaaaca ggaaggcaaa atgcccggaaa	2400
aaagggaata agggcgacac ggaaatgttg aataactcata ctttccctt ttcaatatta	2460
ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa	2520
aaataaaacaa ataggggttc cgccacatt tccccgaaaaa gtgccacctg acgtctaaga	2580
aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc ctttcgttc	2640
gcccgtttcg gtgatgacgg tgaaaacctc tgacacatgc agctccggaa gacggtcaca	2700
gcttgtctgt aagcggatgc cgggagcaga caagccgtc agggcggtc agcgggtgtt	2760
ggcggtgtc gggctggct taactatgcg gcatcagagc agattgtact gagagtgcac	2820
catatgcgt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggcgccat	2880
tcgcccattca ggctgcgcaa ctgttggaa gggcgatcgg tgcgggcctc ttgcgttata	2940
cgccagctgg cgaaaggggg atgtgctgca aggcgattaa gttgggtaac gccagggttt	3000

aac ccc cag ctc tgc tac cag gac acg att ttg tgg aag gac atc ttc	4059
Asn Pro Gln Leu Cys Tyr Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe	
155 160 165	
cac aag aac aac cag ctg gct ctc aca ctg ata gac acc aac cgc tct	4107
His Lys Asn Asn Gln Leu Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser	
170 175 180	
cgg gcc tgc cac ccc tgt tct ccg atg tgt aag ggc tcc cgc tgc tgg	4155
Arg Ala Cys His Pro Cys Ser Pro Met Cys Lys Gly Ser Arg Cys Trp	
185 190 195 200	
gga gag agt tct gag gat tgt cag agc ctg acg cgc act gtc tgt gcc	4203
Gly Glu Ser Ser Glu Asp Cys Gln Ser Leu Thr Arg Thr Val Cys Ala	
205 210 215	
ggt ggc tgt gcc cgc tgc aag ggg cca ctg ccc act gac tgc tgc cat	4251
Gly Gly Cys Ala Arg Cys Lys Gly Pro Leu Pro Thr Asp Cys Cys His	
220 225 230	
gag cag tgt gct gcc ggc tgc acg ggc ccc aag cac tct gac tgc ctg	4299
Glu Gln Cys Ala Ala Gly Cys Thr Gly Pro Lys His Ser Asp Cys Leu	
235 240 245	
gcc tgc ctc cac ttc aac cac agt ggc atc tgt gag ctg cac tgc cca	4347
Ala Cys Leu His Phe Asn His Ser Gly Ile Cys Glu Leu His Cys Pro	
250 255 260	
gcc ctg gtc cag tac atc aaa gct aac tcc aaa ttc atc ggt atc acc	4395
Ala Leu Val Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr	
265 270 275 280	
gag ctg cgg tat aca ttc ggc gcc agc tgt gtg act gcc tgt ccc tac	4443
Glu Leu Arg Tyr Thr Phe Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr	
285 290 295	
aac tac ctt tct acg gac gtg gga tcc tgc acc ctc gtc tgc ccc ctg	4491
Asn Tyr Leu Ser Thr Asp Val Gly Ser Cys Thr Leu Val Cys Pro Leu	
300 305 310	
cac aac caa gag gtg aca gca gag gat gga aca cag cgg tgt gag aag	4539
His Asn Gln Glu Val Thr Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys	
315 320 325	
tgc agc aag ccc tgt gcc cga gtg tgc tat ggt ctg ggc atg gag cac	4587
Cys Ser Lys Pro Cys Ala Arg Val Cys Tyr Gly Leu Gly Met Glu His	
330 335 340	
ttg cga gag gtg agg gca gtt acc agt gcc aat atc cag gag ttt gct	4635
Leu Arg Glu Val Arg Ala Val Thr Ser Ala Asn Ile Gln Glu Phe Ala	
345 350 355 360	
ggc tgc aag aag atc ttt ggg agc ctg gca ttt ctg ccg gag agc ttt	4683
Gly Cys Lys Lys Ile Phe Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe	
365 370 375	
gat ggg gac cca gcc tcc aac act gcc ccg ctc cag cca gag gag ctc	4731
Asp Gly Asp Pro Ala Ser Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu	
380 385 390	

caa gtg ttt gag act ctg gaa gag atc aca ggt tac cta tac atc tca	4779
Gln Val Phe Glu Thr Leu Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser	
395 400 405	
gca tgg ccg gac agc ctg cct gac ctc agc gtc ttc cag aac ctg caa	4827
Ala Trp Pro Asp Ser Leu Pro Asp Leu Ser Val Phe Gln Asn Leu Gln	
410 415 420	
gta atc cgg gga cga att ctg cac aat ggc gcc tac tcg ctg acc ctg	4875
Val Ile Arg Gly Arg Ile Leu His Asn Gly Ala Tyr Ser Leu Thr Leu	
425 430 435 440	
caa ggg ctg ggc atc agc tgg ctg ggg ctg cgc tca ctg agg gaa ctg	4923
Gln Gly Leu Gly Ile Ser Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu	
445 450 455	
ggc agt gga ctg gcc ctc atc cac cat aac acc cac ctc tgc ttc gtg	4971
Gly Ser Gly Leu Ala Leu Ile His His Asn Thr His Leu Cys Phe Val	
460 465 470	
cac acg gtg ccc tgg gac cag ctc ttt cgg aac ccg cac caa gct ctg	5019
His Thr Val Pro Trp Asp Gln Leu Phe Arg Asn Pro His Gln Ala Leu	
475 480 485	
ctc cac act gcc aac ccg cca gag gac gag tgt gtg ggc gag ggc ctg	5067
Leu His Thr Ala Asn Arg Pro Glu Asp Glu Cys Val Gly Glu Gly Leu	
490 495 500	
gcc tgc cac cag ctg tgc gcc cga ggg cac tgc tgg ggt cca ggg ccc	5115
Ala Cys His Gln Leu Cys Ala Arg Gly His Cys Trp Gly Pro Gly Pro	
505 510 515 520	
acc cag tgt gtc aac tgc agc cag ttc ctt cgg ggc cag gag tgc gtg	5163
Thr Gln Cys Val Asn Cys Ser Gln Phe Leu Arg Gly Gln Glu Cys Val	
525 530 535	
gag gaa tgc cga gta ctg cag ggg ctc ccc agg gag tat gtg aat gcc	5211
Glu Glu Cys Arg Val Leu Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala	
540 545 550	
agg cac tgt ttg ccg tgc cac cct gag tgt cag ccc cag aat ggc tca	5259
Arg His Cys Leu Pro Cys His Pro Glu Cys Gln Pro Gln Asn Gly Ser	
555 560 565	
gtg acc tgt ttt gga ccg gag gct gac cag tgt gtg gcc tgt gcc cac	5307
Val Thr Cys Phe Gly Pro Glu Ala Asp Gln Cys Val Ala Cys Ala His	
570 575 580	
tat aag gac cct ccc ttc tgc gtg gcc cgc tgc ccc agc ggt gtg aaa	5355
Tyr Lys Asp Pro Pro Phe Cys Val Ala Arg Cys Pro Ser Gly Val Lys	
585 590 595 600	
cct gac ctc tcc tac atg ccc atc tgg aag ttt cca gat gag gag ggc	5403
Pro Asp Leu Ser Tyr Met Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly	
605 610 615	
gca tgc cag cct tgc ccc atc aac tgc acc cac tcc tgt gtg gac ctg	5451
Ala Cys Gln Pro Cys Pro Ile Asn Cys Thr His Ser Cys Val Asp Leu	
620 625 630	

gat gac aag ggc tgc ccc gcc gag cag aga gcc agc cct ctg acg tcc 5499
 Asp Asp Lys Gly Cys Pro Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser
 635 640 645

ttc aac aac ttc acc acc gtg agc ttc tgg ctg cgc gtg ccc aag gtg agc 5547
 Phe Asn Asn Phe Thr Val Ser Phe Trp Leu Arg Val Pro Lys Val Ser
 650 655 660

gcc agc cac ctg gag atc gtc tct gcg gtg gtt ggc att ctg 5589
 Ala Ser His Leu Glu Ile Val Ser Ala Val Val Gly Ile Leu
 665 670 675

tagaagcttg gtacccgagct cggatccact agtccagtgt ggtggaattc tgcagatatc 5649
 cagcacacgtg gc 5661

<210> 2
 <211> 696
 <212> PRT
 <213> Artificial sequence

<220>
 <223> Recombinant expression plasmid derived from pSI

<400> 2

Met Lys Leu Cys Ile Leu Leu Ala Val Val Ala Phe Val Gly Leu Ser
 -15 -10 -5

Leu Gly Met Lys His Gln His Gln His Gln His Gln His Gln His Gln
 -1 1 5 10

Ala Pro Ser Thr Gln Val Cys Thr Gly Thr Asp Met Lys Leu Arg Leu
 15 20 25 30

Pro Ala Ser Pro Glu Thr His Leu Asp Met Leu Arg His Leu Tyr Gln
 35 40 45

Gly Cys Gln Val Val Gln Gly Asn Leu Glu Leu Thr Tyr Leu Pro Thr
 50 55 60

Asn Ala Ser Leu Ser Phe Leu Gln Asp Ile Gln Glu Val Gln Gly Tyr
 65 70 75

Val Leu Ile Ala His Asn Gln Val Arg Gln Val Pro Leu Gln Arg Leu
 80 85 90

Arg Ile Val Arg Gly Thr Gln Leu Phe Glu Asp Asn Tyr Ala Leu Ala
 95 100 105 110

Val Leu Asp Asn Gly Asp Pro Leu Asn Asn Thr Thr Pro Val Thr Gly
115 120 125

Ala Ser Pro Gly Gly Leu Arg Glu Leu Gln Leu Arg Ser Leu Thr Glu
130 135 140

Ile Leu Lys Gly Gly Val Leu Ile Gln Arg Asn Pro Gln Leu Cys Tyr
145 150 155

Gln Asp Thr Ile Leu Trp Lys Asp Ile Phe His Lys Asn Asn Gln Leu
160 165 170

Ala Leu Thr Leu Ile Asp Thr Asn Arg Ser Arg Ala Cys His Pro Cys
175 180 185 190

Ser Pro Met Cys Lys Gly Ser Arg Cys Trp Gly Glu Ser Ser Glu Asp
195 200 205

Cys Gln Ser Leu Thr Arg Thr Val Cys Ala Gly Gly Cys Ala Arg Cys
210 215 220

Lys Gly Pro Leu Pro Thr Asp Cys Cys His Glu Gln Cys Ala Ala Gly
225 230 235

Cys Thr Gly Pro Lys His Ser Asp Cys Leu Ala Cys Leu His Phe Asn
240 245 250

His Ser Gly Ile Cys Glu Leu His Cys Pro Ala Leu Val Gln Tyr Ile
255 260 265 270

Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Leu Arg Tyr Thr Phe
275 280 285

Gly Ala Ser Cys Val Thr Ala Cys Pro Tyr Asn Tyr Leu Ser Thr Asp
290 295 300

Val Gly Ser Cys Thr Leu Val Cys Pro Leu His Asn Gln Glu Val Thr
305 310 315

Ala Glu Asp Gly Thr Gln Arg Cys Glu Lys Cys Ser Lys Pro Cys Ala
320 325 330

Arg Val Cys Tyr Gly Leu Gly Met Glu His Leu Arg Glu Val Arg Ala
335 340 345 350

Val Thr Ser Ala Asn Ile Gln Glu Phe Ala Gly Cys Lys Lys Ile Phe
355 360 365

Gly Ser Leu Ala Phe Leu Pro Glu Ser Phe Asp Gly Asp Pro Ala Ser
370 375 380

Asn Thr Ala Pro Leu Gln Pro Glu Gln Leu Gln Val Phe Glu Thr Leu
385 390 395

Glu Glu Ile Thr Gly Tyr Leu Tyr Ile Ser Ala Trp Pro Asp Ser Leu
400 405 410

Pro Asp Leu Ser Val Phe Gln Asn Leu Gln Val Ile Arg Gly Arg Ile
415 420 425 430

Leu His Asn Gly Ala Tyr Ser Leu Thr Leu Gln Gly Leu Gly Ile Ser
435 440 445

Trp Leu Gly Leu Arg Ser Leu Arg Glu Leu Gly Ser Gly Leu Ala Leu
450 455 460

Ile His His Asn Thr His Leu Cys Phe Val His Thr Val Pro Trp Asp
465 470 475

Gln Leu Phe Arg Asn Pro His Gln Ala Leu Leu His Thr Ala Asn Arg
480 485 490

Pro Glu Asp Glu Cys Val Gly Glu Gly Leu Ala Cys His Gln Leu Cys
495 500 505 510

Ala Arg Gly His Cys Trp Gly Pro Gly Pro Thr Gln Cys Val Asn Cys
515 520 525

Ser Gln Phe Leu Arg Gly Gln Glu Cys Val Glu Glu Cys Arg Val Leu
530 535 540

Gln Gly Leu Pro Arg Glu Tyr Val Asn Ala Arg His Cys Leu Pro Cys
545 550 555

His Pro Glu Cys Gln Pro Gln Asn Gly Ser Val Thr Cys Phe Gly Pro
560 565 570

Glu Ala Asp Gln Cys Val Ala Cys Ala His Tyr Lys Asp Pro Pro Phe
575 580 585 590

Cys Val Ala Arg Cys Pro Ser Gly Val Lys Pro Asp Leu Ser Tyr Met
595 600 605

Pro Ile Trp Lys Phe Pro Asp Glu Glu Gly Ala Cys Gln Pro Cys Pro
610 615 620

Ile Asn Cys Thr His Ser Cys Val Asp Leu Asp Asp Lys Gly Cys Pro
625 630 635

Ala Glu Gln Arg Ala Ser Pro Leu Thr Ser Phe Asn Asn Phe Thr Val
640 645 650

Ser Phe Trp Leu Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Ile
655 660 665 670

Val Ser Ala Val Val Gly Ile Leu
675